In-vivo identification and control of aerotaxis in Bacillus subtilis

2016 IEEE 55th Conference on Decision and Control (CDC)
ARIA Resort & Casino
December 12-14, 2016, Las Vegas, USA

In-vivo identification and control of aerotaxis in *Bacillus subtilis*

Filippo Menolascina, Roman Stocker and Eduardo D. Sontag

***Abstract*— Problems of identification and control of biological systems have recently attracted an increasing amount of attention. Most work in this field has classically focused either on gene or protein networks. In this manuscript, we focus on the control of a behavioral trait emerging from a signaling network: aerotaxis, the directed motion of bacteria towards (or away from) oxygen. To do so, we consider a bacterium, *Bacillus subtilis*, which is strongly attracted by oxygen, and we quantitatively probe the dynamics of accumulation of populations of this microorganism when exposed to tightly controlled gradients of oxygen generated in a microfluidic device. Combining *in-vivo* experiments with system identification methods, we determine a simple model of aerotaxis in *B. subtilis*, and we subsequently employ this model in order to compute the sequence of oxygen gradients needed to achieve regulation of the center of mass of the bacterial population. We then successfully validate both the model and the control scheme, by showing that *in-vivo* positioning control can be achieved via the application of the precomputed inputs in an open-loop configuration.**

I. INTRODUCTION

Systems and Synthetic Biology play a central role in providing an engineering framework to quantitatively study and synthesize the biomolecular circuits governing cellular functions. At the nexus of these disciplines, control theorists and engineers have developed a number of fundamental results in the past decade demonstrating how elegant theory [8] could be combined with advanced technologies (e.g. microfluidics) [5] to identify [7], [6], [12] and control [9], [10], [14] cellular networks. While most of the past research in this field has focused either on gene [9], [15] or signaling networks [13], comparatively very little effort has been devoted to regulating the often complex behaviors stemming from gene/protein networks, such as cellular motility [3].

To help address this gap, we study here the problem of *in-vivo* control of a behavioral trait in prokaryotes stemming from a known signaling pathway: aerotaxis, the directed migration of bacteria towards (or away from) oxygen. Enabling a key process such as energy harvesting through respiration, it comes as no surprise that many motile bacteria display aerotaxis [2].

This behavior is particularly strong in *Bacillus subtilis*, a widespread bacterium, once used as immunostimulant probiotic for human use [16] and now widely adopted in biotechnology, e.g. as biosensor for food spoilage detection [4]. Given these characteristics (a strong aerotactic response combined with adaptability to heterogeneous environmental conditions) and the potential uses of this microorganism in industrial biotechnology and synthetic biology, we have selected it as model system for our experiments.

In this contribution we combine *in-vivo* experiments with *in-silico* model identification, analysis and control design in order to achieve positioning control of a population of *B. subtilis* exposed to one-dimensional oxygen gradients. In particular, we will be concerned with the control of the first moment of the distribution of a population of bacteria, i.e. their center of mass. To achieve this goal, we take a system identification approach, and expose dilute suspensions of *B. subtilis* to time-varying, opposite oxygen gradients in a microfluidic device. We use measured data in order to extract a dynamical model of how the center of mass of the bacteria varies in response to these inputs. We then analyze the obtained model, and identify several distinctive aspects of it, such as delays. We propose a feedback control scheme based on a Smith Predictor and a relay controller meant to achieve set-point regulation and signal tracking, and test this control algorithm *in-silico*, showing that it is indeed able to achieve a satisfactory performance. Finally, to investigate the performance of both the proposed model and the control scheme, we design and implement an open-loop control experiment in which we apply *in-vivo* the input that had been computed *in-silico* for achieving set-point regulation and signal tracking. The quantitative agreement between the experimental quantifications and the predictions confirms that our approach is indeed able to constrain the center of mass of a population of bacteria so as to display a desired behavior.

The manuscript is organized as follows: Section II presents both the experimental techniques (microfluidics, videomicroscopy, image processing, system identification and control system design) and the design of the experiments used to model and control aerotaxis in *B. subtilis*; Section III presents the *in-silico* results on

This work was supported by National Institutes of Health Grant 1R01GM100473-01 (to R.S. and E.D.S.) and by the Wellcome Trust-University of Edinburgh Institutional Strategic Support Fund (to F.M.)

F. Menolascina is with the Institute for Bioengineering, School of Engineering, University of Edinburgh, The King's Buildings, EH9 3BF, Edinburgh, UK filippo.menolascina@ed.ac.uk

R. Stocker is with the Department of Civil, Environmental and Geomatic Engineering, ETH Zurich, 8063 Zurich, Switzerland romanstocker@ethz.ch

Eduardo Sontag is with the Mathematics Department, Rutgers University, Piscataway, NJ, USA eduardo.sontag@gmail.com

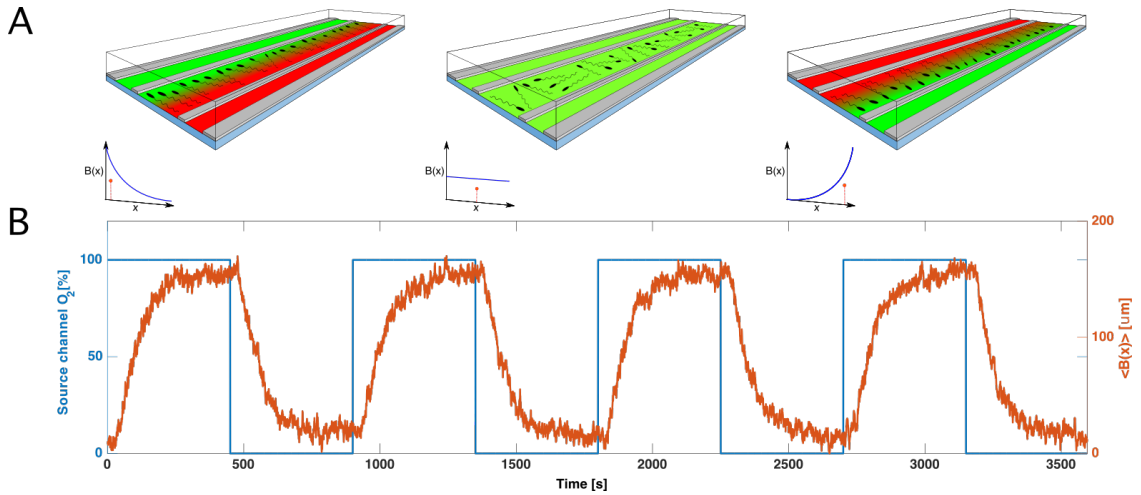

Fig. 1. **Microfluidic setup and design of the identification experiments**. (A) Oxygen, in green, and nitrogen, in red, are flown in the channels on the left (sink) and right (source) of the central channel (test). The PDMS walls (in gray) allow the diffusion of gas species therefore enabling the generation of an oxygen gradient in the test channel. Here bacteria will respond and accumulate towards oxygen rich areas: imaging them and extracting their coordinates allows to compute their spatial distribution $B(x)$. The three panels show the ideal case of a linear negative (left panel), null (center panel) and positive (right panel) oxygen gradient, as well as the related spatial distributions $B(x)$ (in blue in the plots) and their center of mass (in orange). (B) The dynamics of the center of mass (orange), in response to a sequence of oxygen gradients (here summarized by the oxygen concentration in the source channel, see text) is here plotted as a function of time.

model identification and control system test as well as the *in-vivo results*. Finally in Section IV we gather our concluding remarks and future directions for this work.

II. MODEL DERIVATION AND CONTROL SCHEME DESIGN

A. Experimental setup

Here we describe the setup of the experimental platform we used for the identification and control experiments as well as their design.

1) Bacteria, microfluidics and video-microscopy: *B. subtilis* strain OI1085 cells were inoculated in 2 mL of Cap Assay Minimal media and incubated overnight at 37°C, 250 rpm. The culture was allowed to reach 0.3 OD_{600}, then it was diluted 1:10 in fresh media and injected into the microfluidic device (Fig. 1, central channel). This device, composed of three parallel channels 600 μm $\times$ 100 μm $\times$ 2 cm separated by 200 μm walls, was fabricated in oxygen permeable Polydimethylsiloxane (PDMS), using standard replica molding of a silicon-SU8 master. Gas permeability of PDMS allows us to generate oxygen gradients in the test channel by flowing either oxygen/nitrogen in the channel on the left of the central (sink) and nitrogen/oxygen in the opposite (source). The injection of either gases (100% Nitrogen and 100% Oxygen) is controlled by four mass flow controllers (Coler Parmer) driven by a computer via a National Instruments DAQ board. A 20$\times$ objective, in combination with an Andor Zyla camera with 6.5 μm/pixel (0.33 μm/pixel resolution), was used to image an area of the central channel at mid depth (50 μm from the glass slide) and mid length (1 cm). NIS Elements 4.10 was used to acquire 2 frames per second videos of the imaging field for the duration of the experiments. The images in these sequences were then segmented using a custom MATLAB code to identify the x-y coordinates of each cell (see II-B.1). The center of mass of the population was obtained by averaging the x coordinates of the cells over the entire population.

2) Design and execution of the experiments: In order to identify a mathematical model of the dynamics of the center of mass of the population of bacteria, we generated two types of spatial gradients of oxygen: (a) 100%-0% or (b) 0%-100%, oxygen concentration in the sink and source channels respectively. As, according to this choice, the concentration of oxygen in one channel (Ox, expressed as a fraction of oxygen saturation in air) unequivocally determines the concentration of oxygen in the other channel (100%-Ox) we then consider the concentration in the source channel as our input (note that since the source channel is located on the right of the test channel, applying Ox=100%=1 induces bacteria to move towards increasing x and, therefore, pushes the position of their center of mass towards higher values. All the experiments reported in this manuscript have been carried out following the general procedure detailed here: (a) a 100%-0% O_2 gradient was applied for 10 minutes before the start of the experiment ("initialization", meant to stabilize the center of mass of the bacteria towards the minimum feasible value) and (b) then the desired input waveform was applied. For our identification experiments, we applied square waves of

the input of half period equal to 2, 3, 4, 6, 8, 9, 10, 12, 13 and 15 minutes (Fig. 1B shows the case of half period equal to 8 min) for a time between 12 and 60 minutes.

B. System identification, analysis and control

1) Data acquisition and preliminary analysis: Data were acquired in the form of phase contrast micrographs and processed using a custom image segmentation algorithm developed in MATLAB. This procedure was designed to identify single cells and their spatial coordinates in the microfluidic channel. To this aim we used a Canny edge detector that exploits the properties of phase contrast (large intensity variations where two objects with different densities come in contact, water and bacteria in this case). This first step returns a binary image with non-zero elements being the identified edges of cells. Morphological transformations ('closure' and 'fill') and Gaussian filtering then allowed to obtain a binary image where non-zero elements represented pixels belonging to an individual cell. We obtained spatial coordinates of single cells by extracting the centroid for each of these areas (identified as connected components). As we expose cells to a 1D gradient defined over the width of the channel (x axis in Fig. 1A, i.e. $\nabla_y O_2 = 0$), we are only interested in the x-coordinate. For each image we can then plot the spatial distribution of bacteria along the x-axis, $B(x)$ (see Fig. 1A) and extract the center of mass of such distribution $< B(x) >$, defined as:

$$< B(x) > = \int_0^L x \cdot B(x)\, dx - B_{\text{ini}} \tag{1}$$

where $L = 600\ \mu$m is the width of the test channel in our microfluidic device and $B_{\text{ini}} = \min \int_0^L x \cdot B(x)\, dx$ calculated during the initialization step. For each micrograph, and therefore for each time t, we then obtain the center of mass of the bacteria $< B(x,t) >$, i.e. the variable we aim to control. Applying this analysis to the *in-vivo* data described above allowed us to gain some fundamental insights into the dynamics of aerotaxis in *B. subtilis*: (a) a time-delay is present between the input and the output (gradient switch and measured center of mass of the bacteria, compare blue and orange signals in Fig. 1B), (b) the variable we aim to control seems to behave as a low order LTI system. While a formal explanation for (b) might be non-trivial to identify, the origin of the time-delay in the system is very likely to lie in the interplay between the dynamics of oxygen sensing in *B. subtilis* and the diffusion dynamics of oxygen in the test channel (from the approximation of the diffusion equation [1] we desume it takes $t \approx \frac{L}{2D} = \frac{(600\ \mu\text{m})^2}{2*3500\ \mu\text{m}^2/s} = 51\ s$ to establish the oxygen gradient over a 600 μm distance -with $D = 3500\ \mu\text{m}^2/s$ diffusion coefficient of oxygen in water). Note that, although this is the time required to form a fraction >90% of the gradient, bacteria will actually start responding to it long before then. In fact, results from a complementary study we undertook, suggest that *B. subtilis* rescales gradients logarithmically, leading to a significant amplification of its abilities to sense oxygen gradients -the 14 s delay identified by the system identification algorithm should be interpreted in light of this.

2) Model identification and control scheme synthesis: Based on these observations, we set out to identify a mathematical model of the dynamics of $< B(x) >$, the center of mass of a population of aerotaxing *B. subtilis*. In particular, given the observations made in the previous section we seek to identify a time-delayed transfer function. We then have to estimate (a) the magnitude of the time delay and (b) the order and (c) zeros-poles-gain of the system. To this aim we developed a MATLAB script that, given a set of input-output data (i.e. the 10 experiments previously described) automatically estimates the (a) the input-output time delay and (b) the zeros/poles/gain of the transfer function that minimizes the Sum of the Squared Errors between model prediction and experimental data:

$$SSE = \sum_{j=1}^{10} \sum_{t=1}^{T_j} (y_{j,m}(t) - y_{j,e}(t))^2 \tag{2}$$

where $y_{j,m}(t)$ is the predicted output for the jth experiment at time t, $y_{j,e}(t)$ is the experimental quantity for the same combination experiment/time and T_j is number of samples in the jth experiment. An iterative parameter identification routine based on the Gauss-Newton method is used to then identify models of all the feasible combinations of systems of the type:

$$G_d(s) = G(s)\, e^{-\delta s} \tag{3}$$

with

$$G(s) = K \frac{\prod_{l=0}^{2}(s - z_l)}{\prod_{m=0}^{2}(s - p_m)} \tag{4}$$

and K static gain, z_l l^{th} zero, p_m m^{th} pole of the system and $s = j\omega$ complex frequency. To compare the accuracy of the $G_{d,k}$ ($k = 1,..,5$) models we calculate, for each of them, the Normalized Mean Root Square Error averaged over the experimental dataset:

$$< NRMSE_k > = \frac{1}{10} \sum_{j=1}^{10} 1 - \frac{||y_{j,e} - y_{j,m}||}{||y_{j,e} - \frac{1}{T_j} \sum_{t=1}^{T_j} y_{j,e}(t)||} \tag{5}$$

The structure and parameters of the selected model are reported in the next section.

In order to devise a control strategy for $< B(x) >$ we need to review the properties of the system at hand: with respect to this we note that the system, although linear, is affected by a input-output delay. In order to compensate for such delay we design a control scheme based on the combination of a Smith predictor [11] and a relay controller (see Fig. 3) implementing the following piecewise constant function

$$\begin{cases} u = 0 & \text{if } \; e > 0 \\ u = 1 & \text{if } \; e \leq 0 \end{cases} \tag{6}$$

In this configuration the cells (in the block scheme represented as the plant P) are exposed to the binary input u (0 representing the gradient 100%-0% and 1 representing 0%-100%). The controller takes the error signal and converts it into the binary input u. This signal is applied to both the plant P and $G(s)$, the model of the plant $G_d(s)$ where $\delta = 0$. Feeding back the signal $\hat{y}$, allows the controller to work on the "anticipated" output of the system before it is actually generated by the plant. Comparing the actual output $y_e =< B(x) >$ and the predicted one y_m, allows to calculate y_c, a signal representing the contribution of noise and unmodeled dynamics. These components are then summed to $\hat{y}$ to correct the output, y_f, that is eventually compared to the reference (desired) center of mass r (Fig. 3). The performance of this control scheme are presented in the following section.

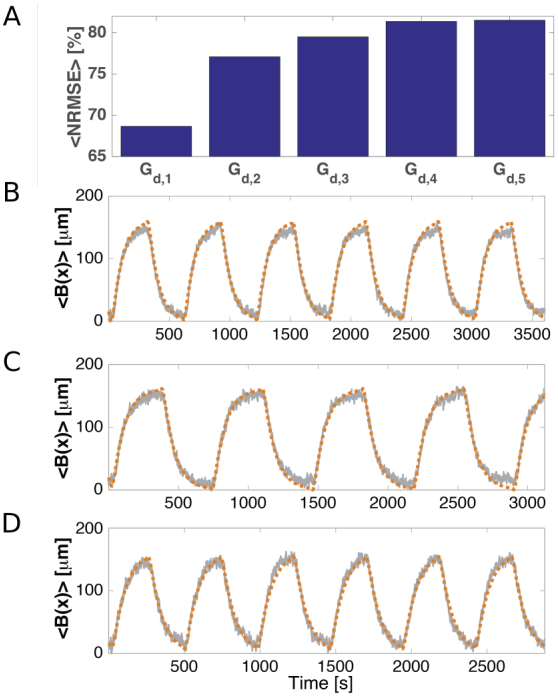

Fig. 2. **System Identification results**. (A) Summary of the <NRMSE> for the models in Table 1. (B)-(D) Experimentally measured center of mass $< B(x) >$, in gray, for three of the 10 experiments in our dataset, plotted with the prediction, dashed orange line, from the selected model, $G_{d,4}$.

TABLE I

MODELS COMPARED IN THIS STUDY

Model	Transfer function
$G_{d,1}$	$\frac{1.598}{s+0.009559}\, e^{-14.5\,s}$
$G_{d,2}$	$-21.329\, \frac{s-0.08399}{s+0.01099}\, e^{-15\,s}$
$G_{d,3}$	$\frac{0.11684}{(s+0.05791)\,(s+0.01244)}\, e^{-14\,s}$
$G_{d,4}$	$-1.0146\, \frac{(s-0.1667)}{(s+0.08472)(s+0.01235)}\, e^{-14\,s}$
$G_{d,5}$	$5.3009\, \frac{s^2-0.2605s+0.03736}{(s+0.09933)(s+0.01233)}\, e^{-15\,s}$

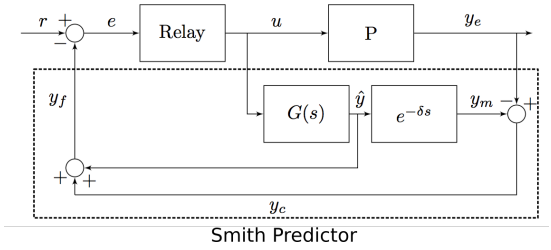

Fig. 3. **Control Scheme**. The proposed control algorithm combines a Smith predictor (dashed box) with a relay controller.

III. RESULTS

A. Areotaxis is well captured by a low order time-delayed LTI system

Using the approach described in the previous section we identified 5 time-delayed transfer functions ($G_{d,k}(s)$ with $k = 1,..,5$, see Table 1 for more details). Their $< NRMSE >$ is presented in Fig. 2A.

While $G_{d,4}$ and $G_{d,5}$ perform significantly better than the others (see Fig. 2A), we choose to use $G_{d,4}$ in the following (see agreement between model and experiments in Fig. 2B-D) since a minimal increase in model accuracy (0.2%) in $G_{d,5}$ comes at the expense of a more complex model (additional zero, see Table 1). Therefore the dynamics of the center of mass of *B. subtilis* during aerotaxis can be captured with high accuracy ($< NRMSE >= 81\%$) with a second order time-delayed linear time invariant (LTI) model (see Fig. 2B-D for a comparison between experimental data in grey and model prediction in orange).

B. In-silico control experiments

We then used this model to simulate the control scheme described in the section II-B.2; in particular we substituted $G(s) = -1.0146 \frac{(s-0.1667)}{(s+0.08472)(s+0.01235)}$ in the block scheme and put $\delta = 14$. We then set out to assess the capabilities of the proposed controller in a standard set-point control task. To this aim we implemented the control scheme in Fig. 3 in MATLAB Simulink, we set $r = 150\,\mu$m and ran a simulation over a 15 min time-horizon ($t = 0,..,900$ s). The results of this simulation is presented in Fig. 4A-B (system output in blue, set-point in black). The control scheme devised here does indeed prove to be effective in forcing the center of mass of the population to oscillate within a tight region around the set-point. Therefore the proposed control scheme appears to be able to effectively constrain the dynamics of the center of mass of the bacterial population, *in-silico*. Would this result hold also *in-vivo*? To test this we decided to apply the input computed in this *in-silico* experiment directly *in-vivo*, to a population of bacteria swimming in a microfluidic device. We describe the results of this experiment in the following section.

C. In-vivo control experiments

We finally set out to achieve positioning control *in-vivo*. We will do so using an "open-loop" configuration: we will apply a pre-computed input (see Fig. 4A, green signal) to generate a sequence of microscale oxygen gradients meant to force the center of mass of a population of cells to reach a desired position, importantly, without any real-time feedback from actual positions of the cells. Being able to accomplish such task would, at the same time, provide a strong indication of both (a) the accuracy of the identified model and (b) the robustness of the proposed control scheme. The results we obtained carrying out this experiment *in-vivo* are presented in Fig. 4A. We note that, besides the expected noise, the measured $< B(x) >$ (red signal) matches very closely the output of the system we simulated in our *in-silico* set-point control experiment. As a matter of fact the agreement between experimental data and predictions is so high that amplitude and frequency of the small oscillations we observe in the measured output for $t > 100$ s match perfectly and, what is more, appear phase-locked until the end of the experiment with the model-predicted ones. As these results seemed to hint at the opportunity to extend the control experiment to longer time-horizons we tested the ability to control $< B(x) >$ over a 1-hour time window. Moreover we decided to proceed to increase the complexity of such experiment by posing it as a signal-tracking control problem. Similarly to what has been done with the setpoint experiment, we carried out the experiment *in-silico*, extracted the control input and applied it *in-vivo*. The results of this experiment, reported in Fig. 4B confirm that the proposed approach is in fact able to force the center of mass of the population of bacteria to follow a time-varying signal rather than just a steady one. It should be noted that some of the limitations of this control strategy (e.g. the oscillations around 2500 min) are due to the control algorithm used; more advanced control schemes (e.g. MPC) can be designed to address such issues.
Taken together these results confirm that our control scheme (a) is in fact robust to noise and (b) is based on a predictive model that does indeed capture the dynamics of *B. subilis*' response to oxygen.

IV. CONCLUSIONS AND FUTURE WORK

In this work we used a combination of *in-silico* modelling and *in-vivo* experiments to derive a simple time-delayed LTI model of the dynamics of the center of mass of a population of *B. subtilis* during aerotaxis. We then used the model we developed to achieve open-loop *in-vivo* positioning control of bacteria in a microfluidic device. A closed-loop implementation of the proposed scheme is currently being developed: however, a few technological challenges (e.g. real-time image processing to extract the center of mass of the population from micrographs) have to be addressed before such a result can be achieved. Nevertheless, the approach we propose proved to be able to successfully accomplish this task and we believe has the potential to enable the exploitation of endogenous motility in bacteria to precisely guide genetically engineered microorganisms towards a target (e.g. the site of a contamination) to accomplish a task (e.g. secrete a chemical species that mitigates the effect of the contamination) and to effectively bring them back in a tightly controlled way. However more research will be needed to make this happen: for example, to achieve a tighter control over the population, real-time control as well as control over moments of the spatial distribution above the 1^{st} will have to be tackled and 2 and 3 dimensional positioning proved.

REFERENCES

[1] Peter Atkins and Julio De Paula. *Elements of physical chemistry*. Oxford University Press, USA, 2013.

[2] R Barak, I Nur, Y Okon, and Y Henis. Aerotactic response of Azospirillum brasilense. *Journal of Bacteriology*, 152(2):643–649, nov 1982.

[3] U Brandt-Pollmann, D Lebiedz, Moritz Diehl, S Sager, and J Schlöder. Real-time nonlinear feedback control of pattern formation in (bio) chemical reaction-diffusion processes: A model study. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 15(3):033901, 2005.

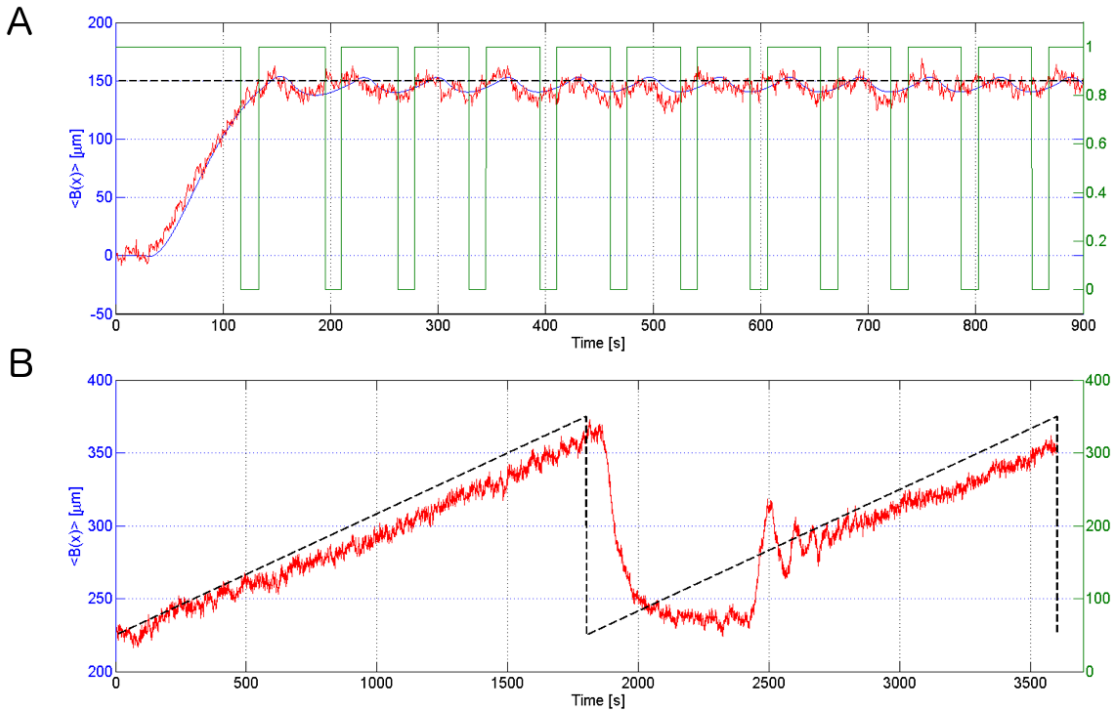

Fig. 4. ***In-silico* and *In-vivo* results**. (A) Set-point and (B) signal-tracking control results on a population of bacteria exposed to pre-computed sequences of oxygen gradients. Set-point in black, model prediction in blue, experimental quantification in red and oxygen concentration in the source channel in green.

[4] Alicja Daszczuk, Yonathan Dessalegne, Isma??l Drenth, Elbrich Hendriks, Emeraldo Jo, Tom Van Lente, Arjan Oldebesten, Jonathon Parrish, Wlada Poljakova, Annisa A. Purwanto, Renske Van Raaphorst, Mirjam Boonstra, Auke Van Heel, Martijn Herber, Sjoerd Van Der Meulen, Jeroen Siebring, Robin A. Sorg, Matthias Heinemann, Oscar P. Kuipers, and Jan Willem Veening. Bacillus subtilis biosensor engineered to assess meat spoilage, 2014.

[5] MS Ferry, IA Razinkov, and J Hasty. Microfluidics for synthetic biology from design to execution. *Methods Enzymol*, 497:295, 2011.

[6] Gianfranco Fiore, Filippo Menolascina, Mario di Bernardo, and Diego di Bernardo. An experimental approach to identify dynamical models of transcriptional regulation in living cells. *Chaos*, 23(2), 2013.

[7] Filippo Menolascina, Domenico Bellomo, Thomas Maiwald, Vitoantonio Bevilacqua, Caterina Ciminelli, Angelo Paradiso, and Stefania Tommasi. Developing optimal input design strategies in cancer systems biology with applications to microfluidic device engineering. *BMC Bioinformatics*, 10 Suppl 12:S4, 2009.

[8] Filippo Menolascina, Mario di Bernardo, and Diego di Bernardo. Analysis, design and implementation of a novel scheme for in-vivo control of synthetic gene regulatory networks. *Automatica*, 47(6):1265–1270, June 2011.

[9] Filippo Menolascina, Gianfranco Fiore, Emanuele Orabona, Luca De Stefano, Mike Ferry, Jeff Hasty, Mario di Bernardo, and Diego di Bernardo. In-Vivo Real-Time Control of Protein Expression from Endogenous and Synthetic Gene Networks. *PLoS Computational Biology*, 10(5), 2014.

[10] Andreas Milias-Argeitis, Sean Summers, Jacob Stewart-Ornstein, Ignacio Zuleta, David Pincus, Hana El-Samad, Mustafa Khammash, and John Lygeros. In silico feedback for in vivo regulation of a gene expression circuit. *Nature biotechnology*, 29(12):1114–6, 2011.

[11] Julio E. Normey-Rico and Eduardo F. Camacho. Dead-time compensators: A survey. *Control Engineering Practice*, 16(4):407–428, 2008.

[12] Jakob Ruess, Francesca Parise, Andreas Milias-Argeitis, Mustafa Khammash, and John Lygeros. Iterative experiment design guides the characterization of a light-inducible gene expression circuit. *Proceedings of the National Academy of Sciences*, 112(26):201423947, 2015.

[13] Jared E Toettcher, Delquin Gong, Wendell A Lim, and Orion D Weiner. Light-based feedback for controlling intracellular signaling dynamics. *Nature methods*, 8(10):837–839, 2011.

[14] Jannis Uhlendorf, Agnès Miermont, Thierry Delaveau, Gilles Charvin, François Fages, Samuel Bottani, Gregory Batt, and Pascal Hersen. Long-term model predictive control of gene expression at the population and single-cell levels. *Proceedings of the National Academy of Sciences of the United States of America*, 109(35):14271–6, 2012.

[15] Jannis Uhlendorf, Agnès Miermont, Thierry Delaveau, Gilles Charvin, François Fages, Samuel Bottani, Gregory Batt, and Pascal Hersen. Long-term model predictive control of gene expression at the population and single-cell levels. *Proceedings of the National Academy of Sciences*, 109(35):14271–14276, 2012.

[16] L. S. Wong, M. S. Johnson, I. B. Zhulin, and B. L. Taylor. Role of methylation in aerotaxis in Bacillus subtilis. *Journal of Bacteriology*, 177(14):3985–3991, 1995.